DOCUMENT RESUME

ED 423 783 HE 031 631

AUTHOR Oppenheim, Nancy
TITLE Undergraduates' Assessment of International Teaching Assistants' Communicative Competence.
PUB DATE 1998-03-18
NOTE 22p.; Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (Seattle, WA, March 1998).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Communicative Competence (Languages); *English (Second Language); *Foreign Students; Higher Education; Screening Tests; *Student Evaluation of Teacher Performance; *Teacher Effectiveness; *Teaching Assistants; Test Validity; Undergraduate Students; Undergraduate Study
IDENTIFIERS *Oral English Language Proficiency Tests

ABSTRACT

This study examined the communicative competence of 89 international teaching assistants (ITAs) based on evaluations by undergraduate students. The study was developed to answer two questions: (1) whether the Oral English Proficiency Test, administered by professionals in teaching English as a second language, is a valid instrument for screening potential ITAs; and (2) the extent to which ITAs' teaching effectiveness improved after attending a mandatory teaching seminar. To address the first question, undergraduates' assessments of ITAs' communicative competence was correlated with ITAs' scores on the Oral English Proficiency Test. To address the second question, ITAs were videotaped teaching lessons to undergraduates, both before and after the teaching seminar, and these tapes were evaluated by a panel of undergraduates. The study found a weak positive relationship between performance on the test and student evaluation of teaching effectiveness and linguistic skill. Undergraduates' evaluations of teaching effectiveness were significantly more positive when they viewed videotapes of ITAs' teaching after the ITA teaching seminar. Results suggest undergraduates' rating of actual teaching videotapes might be a more valid method of screening potential ITAs than the current English screening test. (Contains 16 references.) (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Nancy Oppenheim J.D., M.A., Ph.D
Do not quote from this manuscript
May 31 Draft ©1997

Undergraduates' Assessment of International Teaching Assistants' Communicative Competence

ED 423 783

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Nancy Oppenheim
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Undergraduates' Assessment of International Teaching Assistants' Communicative Competence

Because of the widespread perception that international teaching assistants hindered undergraduate learning at U.S. universities, during the past ten years, Texas and 37 other states mandated oral English language competency standards and/or short courses to insure instructors would be communicatively competent (Oppenheim, 1997). For the purposes of this summative program evaluation research, international teaching assistants (ITAs) were graduate students who were not U.S. citizens or permanent residents and who did not speak English as their native language. The functions of ITAs at this large U.S. research university ranged from lecturing, leading discussions sessions, supervising labs to holding office hours.

Notwithstanding such statutes, university policies and procedures aimed at insuring ITA teaching effectiveness, Fox (1991) at Purdue, Plakans (1994) at Iowa State, and Oppenheim (1997) at The University of Texas at Austin (UT) found that about 50 percent of their representative samples of undergraduates expressed negative attitudes about ITA teaching effectiveness. These findings raised the issue of whether the current assessments of oral English proficiency accurately certified which ITAs would be effective teachers for undergraduates.

Similarly, the large proportion of students' expressing negative evaluations of their ITAs' teaching effectiveness at UT raised the issue of whether the training that ITAs were receiving was improving ITA teaching

effectiveness significantly. The mandatory teaching seminar for ITAs at UT was offered five times per year and consisted of 24 hours of lectures, workshops, and role plays that highlighted how the educational activities, styles of teaching, goals, and classroom climate at UT fostered active student learning. Workshops included "Leading Discussion and Review", "Holding Office Hours/Tutoring", "Interactive Teaching", and "Supervising Labs".

Research Questions

This program evaluation research was undertaken to answer two important questions. First, was the Oral English Proficiency Test administered by professionals trained in the field of Teaching English As A Second Language a valid instrument for screening potential ITAs? In order to explore this question, undergraduates' assessments of ITAs' communicative competence was correlated with ITAs' scores on the Oral English Proficiency Test that was used to certify that ITAs were proficient in Oral English and could be appointed for teaching positions with student contact.

The second research question was the extent to which ITAs' teaching effectiveness improved after attending the mandatory teaching seminar. To examine this question, ITAs were videotaped teaching lessons suitable for introductory classes in their discipline to groups of undergraduates both before and at the conclusion of the teaching seminar. Next, a panel of six undergraduates' evaluated the ITAs' videotapes. The panel's assessments

were used to measure the extent to which the teaching seminar caused noticeable improvement in ITA teaching effectiveness.

Method

Subjects

Eighty-nine prospective ITAs who completed the Oral English Proficiency Exam and the teaching seminar were videotaped and assessed by the undergraduate panel. The ITAs' countries of origin included the People's Republic of China, Taiwan, Hong Kong, Korea, India, Pakistan, Russia, Bulgaria, Germany, Spain, France, Israel, Mexico, and Nigeria. The ITAs' academic disciplines included accounting, aerospace engineering, astronomy, biology, botany, business, chemical engineering, chemistry, civil engineering, computer science, economics, educational psychology, electrical engineering, geology, mathematics, mechanical engineering, nutrition, nursing, pharmacy, physics, social work, and sociology.

Procedure

In order to measure the validity of the current screening instrument, the Oral English Proficiency Test, undergraduates' assessments of ITA teaching effectiveness was correlated with ESL professionals' assessment of ITAs' English proficiency. The undergraduates assessed the videotapes of ITAs teaching a lesson suitable for an introductory class in their discipline to a group of undergraduates. Six undergraduates were employed to evaluate the ITAs' teaching effectiveness because they were representative of the stakeholders who were taught by ITAs in actual discussion and lab sections.

The undergraduates were in different class years and were majoring in different disciplines to mirror the diverse undergraduates taught by ITAs. Survey research revealed that undergraduates from all majors in different class years were taught by ITAs at UT (Oppenheim, 1996). Because the UT student population included a very high proportion of transfer students, many of these students were enrolled in introductory courses as juniors, or seniors. Most students reported being taught by ITAs in introductory courses.

The panel who judged the videotapes consisted of three females: A sophomore majoring in Biology; an Honors Program junior majoring in computer science; and a sophomore majoring in accounting. The three males were: a freshman majoring in humanities; a freshman majoring in petroleum engineering; and a sophomore majoring in liberal arts. All of the undergraduates were native English-speakers. Half of the panel consisted of students who had at least one parent who was a nonnative English-speaker because 50% of the undergraduates at UT reported having at least one parent who was a nonnative English-speaker (Oppenheim, 1997).

The undergraduates went through a two-hour training program to learn how to evaluate the videos using the teaching effectiveness scale. The undergraduates' assessments of videotapes filmed before and after the teaching seminar were randomly arranged to measure whether undergraduates perceived improvement in ITAs' teaching effectiveness after the ITAs completed the mandatory teaching seminar. The undergraduates were told that each ITA was videotaped twice to generate sufficient samples of

teaching behaviors to insure accurate assessment. Undergraduates were only permitted to evaluate the tapes for a maximum of two hours at one sitting to avoid the "fatigue factor" that was noticed by Hinofotis, Bailey and Stern (1981).

Instrumentation

The undergraduate rating form was developed using the research of other ITA developers who sought to construct valid and reliable performance tests of ITAs' communicative competence (Briggs, 1994; Hinofotis & Bailey, 1981; Hinofotis, Bailey, & Stern, 1981). The teaching effectiveness scale consisted of 19 items measuring undergraduates' perceptions of the ITA on four dimensions: Pedagogy, Interpersonal Skills, Linguistic Characteristics and the Undergraduate's familiarity with the Concepts Covered. Each item was judged on a five point scale. The Cronbach alpha reliability coefficient of the Teaching Effectiveness Scale was .91. The Cronbach alpha reliability coefficient of the pedagogical, interpersonal, and linguistic subscales equaled .73, .87 and .85 respectively.

The factor analysis was completed using a varimax rotated extraction of principal components. The factor analysis confirmed that the teaching effectiveness scale measured the four constructs that it was designed to measure. The combined factors explained 65.3% of the variance on the 19 item scale. Sample items from each of the four dimensions and their loadings on the factor are set forth in Table 1.

BEST COPY AVAILABLE

Table 1

Teaching Effectiveness Scale Factors and Sample Item Loadings

Factor	Sample Item	Loading
Interpersonal	The ITA smiled frequently.	.77
	The ITA encouraged students.	.74
	The ITA created a warm and friendly classroom environment.	.79
	The ITA was too formal. (reverse scored)	.75
Linguistic	The ITA raised the pitch of his/her voice to stress important words.	.82
	The ITA's speech was easy to understand	.81
Pedagogy	The ITA provided a preview of what would be covered in the lesson.	.67
	The ITA used a sufficient number of illustrative examples.	.75
Student's Familiarity with Concepts	I was already familiar with the concepts that the ITA spoke about.	.73
		.73

The four factors that explained the greatest proportion of unique variance measured by the ITA teaching effectiveness scale were Interpersonal Skills (38.7%), Pedagogical Skills (13.2%), Linguistic Characteristics (6.8%) and Student's Familiarity with Concepts (5.3%).

The items that were most highly correlated with the ITAs' scores on the ITA teaching effectiveness scale are set forth in Table 2.

Table 2

Item-Total Correlations

Item	Correlation
I would like to have this person as my TA.	.75
The TA created a warm and friendly classroom environment.	.74
The ITA encouraged students.	.71
The ITA's speech was easy to understand.	.68
The ITA was too formal. (reverse scored)	.68
The ITA spoke fluently by paraphrasing with different vocabulary words.	.64
The ITA smiled frequently.	.62

Results

Intra-rater reliability

The first concern was to establish intra-rater reliability. The undergraduates were asked to evaluate ten of the ITAs' videos twice. Their first and second viewing occurred five weeks apart. During the first and second viewings of the ten ITAs, their order was altered. The Pearson product moment correlation coefficient for the undergraduates ranged from .80 to .95. ($p<.01$). All of the raters were significantly consistent in assessing the ITAs' teaching effectiveness. Using a paired sample F-test, the undergraduates' evaluation of the ITAs on their first and second viewing were not significantly different.

Whereas intra-rater reliability was found to be high, inter-rater reliability was expected to vary widely. One of the dimensions built into the teaching effectiveness scale was the undergraduates' familiarity with the concepts taught during the ITAs' lessons. The teaching effectiveness scale was designed to capture the undergraduates' better ability to construct meaning where they possessed prior knowledge of the concepts taught in the lesson. Accordingly, since the undergraduates on the evaluation panel were chosen because of their differing majors and differing conceptual knowledge, it was expected that these differences would be directly related to their differing evaluations of the ITAs.

The inter-rater reliability for the panel as measured by the coefficient alpha was .61. This correlation represents moderately strong agreement among the undergraduates as to the ITAs' teaching effectiveness. The correlation between the undergraduates' responses to item 23 "I was already familiar with the concepts that the ITA spoke about" and the teaching effectiveness scale score was significantly positive ($r=.29$ $p<.001$). Students who had more familiarity with the concepts that the ITAs were teaching tended to evaluate the ITAs more positively as was expected.

Research Question 1

Was the Oral English Proficiency Test that was scored by ESL professionals valid for the purpose of screening potential ITAs?

In response to the first research question about the extent to which the undergraduates' evaluation of the ITAs' teaching effectiveness was correlated

BEST COPY AVAILABLE

to the ESL professionals' evaluation of the ITAs' oral English proficiency, there was a weak positive relationship. The Pearson product moment correlation between the Oral English Proficiency Test score and the teaching effectiveness scale was .25. There was also a weak positive relationship between the ITAs' score on the Oral English Proficiency Test and the undergraduates' evaluations of the ITAs' linguistic skills. The Pearson product moment correlation between the Oral English Proficiency Test score and the linguistic subscale of the teaching effectiveness scale was .33. This finding is similar to the findings of prior studies comparing ESL raters and undergraduate raters of ITAs. The Spearman rank order correlation measured by Hinofotis and Bailey (1981) was .62. In sum, the evaluations of undergraduates about ITAs' oral English proficiency and ESL professionals' assessments diverged significantly.

Research Question 2

To what extent did the ITAs' teaching effectiveness improve after attending the mandatory teaching seminar?

Paired sample F-tests were conducted on the teaching effectiveness scores on the pre- and post-videotapes. 89 ITAs were judged on the 19 item teaching effectiveness scale. The undergraduates expressed the extent of their agreement with each item such as "The ITA encouraged students". A rating of "1" expressed strong disagreement with the item and a "5" expressed strong agreement with the item. Negatively worded items were reverse scored. The average of the 19 items equaled the ITA's teaching effectiveness

score. The mean score on the pre-teaching seminar tapes was 3.13 (s.d. =.86). The mean score on the post-teaching seminar tapes was 3.50 (s.d. =.78). The range of scores on the pre-tapes was a low of 1.21 and a high of 4.84. The range of scores on the post-tapes was a low of 1.74 and a high of 4.95. The one way analysis of variance reflected a significant improvement between the students' evaluations of the ITAs' pre and post tapes ($F = 29.47$ $p < .001$).

The undergraduates' evaluations of the ITAs' interpersonal, pedagogical, and linguistic behaviors were significantly more positive on the post-videotapes. Table 3 sets forth pre- and post- rating means for each item that the undergraduates evaluated significantly higher on the post-videotapes.

Table 3

Pre- and Post- Item Means and F-tests

Item	Pre-tape Mean	Post-tape Mean	F	p value<
Interpersonal				
The ITA turned his/her back to the students too much. (reverse scored)	4.08	4.51	21.11	.001
The ITA smiled frequently.	2.56	2.94	10.12	.002
The ITA made appropriate eye contact.	3.77	4.28	19.67	.001
The ITA encouraged students.	2.68	3.23	23.94	.001
I would like to have this person as my TA.	2.85	3.14	6.54	.01
Pedagogy				
The ITA provided a preview of what would be covered in the lesson.	3.87	4.10	6.42	.01
The ITA used a sufficient number of illustrative examples.	3.44	3.96	25.75	.001
The examples were too advanced or complex to be helpful.	3.92	4.28	14.79	.001
The ITA used helpful diagrams or visuals.	3.24	3.77	22.74	.001
The ITA asked the students questions throughout the lesson.	1.99	3.03	70.62	.001
The ITA provided an outline or summary of the concepts covered.	2.90	3.17	4.37	.037
Linguistic				
The ITA raised the pitch of his/her voice to stress important words.	3.44	3.87	17.69	.001
The ITA spoke fluently by paraphrasing with different vocabulary words.	3.03	3.53	22.28	.001

The family-wide error rate <.05.

Table 3 reflects that undergraduates perceived improvement across all three dimensions of the teaching effectiveness scale after the ITAs completed the teaching seminar.

Regression Analysis Across Constructs and Discrete Variables

All three constructs explained a unique portion of the variance measured by the teaching effectiveness scale. The Interpersonal subscale explained 48% of the variance. The Pedagogy subscale explained 33% and the Linguistic subscale explained 19% of the variance.

To help determine which items explained the most variance in the scale, stepwise regression analyses were run on the data. The seven items that explained the most variance (R square & Adjusted R square=.91) were:

Interpersonal
The ITA smiled frequently.
The ITA encouraged students.
I would like to have this person as my TA.
Linguistic
The ITA's speech was easy to understand.
Pedagogy
The ITA provided a preview of what would be covered in the lesson.
The examples were too advanced or too complex to be helpful. (reverse scored)
The ITA provided an outline or a summary of the concepts covered.

Interestingly, the linguistic items explained very little of the variance measured by the teaching effectiveness scale.

Discussion

Validity of the Oral English Proficiency Test

The results of this program evaluation research revealed that the undergraduates' perceptions of the ITAs' teaching effectiveness was only weakly correlated to the ITAs' scores on the Oral English Proficiency Test. The Oral English Proficiency test focused on the ITAs' pronunciation (30%),

grammar (20%), fluency (10%), and comprehensibility (40%) when they summarized an article from their discipline, pronounced terms, explained two terms, interpreted a graph and responded to questions.

Because the Oral English Proficiency Test was not an accurate predictor of the students' evaluations of the ITAs' teaching effectiveness, it appears that the Oral English Proficiency Test fails to measure the appropriate constructs for predicting whether an ITA would be able to teach undergraduates effectively. Briggs (1994) and Bailey (1985) researched the features that an oral performance test for ITAs should cover. It was suggested that strategic and sociolinguistic competence was more important than linguistic factors. Similarly, Hoejke and Williams, (1994) proposed that the goal of ITA education should be communicative competence in the classroom rather than concentrating narrowly on linguistic skills. They suggested that developing pragmatic competence and compensatory strategies would improve ITA teaching effectiveness.

From the research on effective college teaching, students evaluated teachers who were willing to help students, and who were organized most favorably (Briggs & Hofer, 1992; Civikly, 1992; Feldman, 1989; Marsh, 1984; McKeachie, 1994). By contrast, The Oral English Proficiency Test focused narrowly on linguistic skills rather than also testing pedagogical, and interpersonal skills.

The Oral English Proficiency Test needed to be more accurate because even though all of the ITAs achieved a score of at least 230 out of 300 or 77%

BEST COPY AVAILABLE

on the Test, about fifty percent of the undergraduates currently being taught by ITAs evaluated the quality of ITA instruction as less than satisfactory (Oppenheim, 1997). The Oral English Proficiency Test did not appear to accurately predict which ITAs would be able to meet the legislative standard that "all courses be taught clearly in English" (Texas Revised Statutes, 1987, 1989). The statute was enacted to insure that undergraduates received appropriate instruction.

It is suggested that a panel of undergraduates could rate videotapes of ITAs teaching introductory lessons using the teaching effectiveness scale to accurately screen prospective ITAs' communicative competence for teaching appointments.

Monetarily, the current Oral English Proficiency Test costs $50.00 for each prospective ITA. For each assessment, two ESL raters were paid $50.00 per hour. On average, they completed assessments of two prospective ITAs per hour. By comparison, it would cost $11.60 for each prospective ITA to be videotaped and assessed by a panel of six undergraduates. Four undergraduates would be hired at $7.00 per hour to be taught by ITAs and videotaped. About seven prospective ITAs could be videotaped per hour. One graduate student would videotape the micro-teaching session and be paid $11.00 per hour. A panel of six undergraduates would be hired at $7.00 per hour to rate the videotapes. About seven prospective ITAs' videotapes could be assessed per hour.

Before teaching effectiveness scale could be put in place as a screening device, cut-scores for different types of teaching appointments in different departments would have to be set. There needs to be strict scrutiny of the validity of any screening device and cut-score established that is used to determine whether ITAs can be employed to teach undergraduates. No ITA should be barred from earning a livelihood arbitrarily.

The Effectiveness of the Mandatory Teaching Seminar

The teaching seminar was mandated by House Bill 638 which required that all public universities in Texas provide a program or short course to ensure that all courses are taught clearly in English (Texas Revised Statutes, 1989). The teaching seminar caused significant improvement in undergraduates' perceptions of ITAs' teaching effectiveness. The undergraduates who evaluated the pre-and post-videotapes scored the ITAs' post-videotapes significantly higher on interpersonal, pedagogical and linguistic skills. Interpersonally, the undergraduates noted significant improvement in the ITAs' use of eye contact, smiling, and encouragement of students. Pedagogically, the undergraduates noted significant improvement in the ITAs' previews and summaries of what would be covered in the lesson, asking students questions throughout the lesson, and the use of helpful examples, diagrams and visuals. Linguistically, the undergraduates evaluated the ITAs' post-videotapes significantly higher on the ITAs' ability to raise the pitch of their voices to stress important words. The ITAs' teaching

effectiveness was also evaluated significantly higher because they spoke more fluently by paraphrasing with different vocabulary words.

The greatest variance in teaching effectiveness scores was explained by the ITAs' interpersonal and pedagogical skills. The undergraduates' assessments of the ITAs' pre- and post-videotapes reflected significant improvement in crucial dimensions of teaching effectiveness.

An important question that remains unanswered is whether the ITAs' improvement in these skills will be maintained once they are on the job. Backsliding is of great concern. There are no targeted professional development services for the ITAs who complete the teaching seminar. A needs analysis of the ITAs' professional development requirements could be undertaken to provide directions for providing appropriate in-service support. The undergraduates' feedback about their ITAs' interpersonal, linguistic and pedagogical skills would be invaluable.

Conclusion

The results of this program evaluation research revealed serious questions as to the validity of the Oral English Proficiency Test to screen prospective ITAs. An alternative screening procedure using undergraduate panels to evaluate prospective ITAs microteaching was proposed.

The ITAs' teaching effectiveness was significantly enhanced by attending the mandatory teaching seminar as judged by undergraduates' evaluations of pre and post-videotapes of the ITAs teaching. The undergraduates consistently evaluated those ITAs with strong interpersonal,

BEST COPY AVAILABLE

pedagogical and linguistic skills as effective teachers. Interpersonal and pedagogical skills were most closely associated with being an effective ITA.

REFERENCES

Bailey, K.M. (1985). If I had known then what I know now: Performance testing of foreign teaching assistants. In P. Hauptman, R. Leblanc, & M. Wesche (Eds.), Second language performance testing (pp. 153-180). Ottawa: University of Ottawa Press.

Briggs, S. (1994). Using performance assessment methods to screen ITAs. In C. Madden, & C. Myers (Eds.), Discourse and performance of international teaching assistants (pp. 63-80). Alexandria, Virginia: TESOL.

Briggs, S., & Hofer, B. (1991). Undergraduate perceptions of ITA Effectiveness. In J. Nyquist, R. Abbot, D. Wulff, & J. Sprague (Eds.), Preparing the professoriate of tomorrow to teach (pp. 435-445). Dubuque, Iowa: Kendall/Hunt Publishing Company.

Civikly, J. (1992). Clarity: Teachers and students making sense of instruction. Communication Education, 41 138-152.

Feldman, K. (1989a). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. Research in Higher Education, 30,6, 583-645.

Fox, W. (1992). Functions and effects of international teaching assistants at a major research institution. (Doctoral dissertation, Purdue University, 1991). Dissertation Abstracts International, 52, 3193-A.

BEST COPY AVAILABLE

Hinofotis, F.B., & Bailey, K. (1981). American undergraduates' reactions to the communication skills of foreign teaching assistants. In J. Fisher, M. Clarke, & J. Schacter (Eds.), On TESOL '80 Building bridges: Research and practice in teaching English as a second language (pp. 120-136). Washington, DC: TESOL.

Hinofotis, F.B., Bailey, K., & Stern, S. (1981). Assessing the oral proficiency of prospective teaching assistants: Instrument development. In A.S. Palmer, P.J. M. Groot & G.A. Trosper (Eds.) The construct validation of tests of communicative competence (pp. 106-126). Washington, DC: TESOL.

Hoekje, B. & Williams, J. (1994). Communicative competence as a theoretical framework for ITA education. In C. Madden, & C. Myers (Eds.), Discourse and performance of international teaching assistants (pp. 11-26). Alexandria, Virginia: TESOL.

Marsh, H. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76.5, 707-754

McKeachie, W. (1994). Teaching tips. Lexington, MA: D.C. Heath & Co.

Oppenheim, N. (1997). "Living through" an intercultural experience: Undergraduates learning from nonnative English-speaking teaching assistants. (Doctoral dissertation, The University of Texas at Austin, 1996). Dissertation Abstracts International, *58-01B* 9719449.

Oppenheim, N. (1997). Law informing policy: Employment of international faculty . Paper presented at the annual meeting of TESOL, Orlando, Florida.

Plakans, B. (1994). Undergraduate experiences with and attitudes toward international teaching assistants. (Doctoral dissertation, Iowa State University, 1994).

Texas Revised Statutes (1987). Article III, Section 28.

Texas Revised Statutes (1989). House Bill 638.

BEST COPY AVAILABLE